Example: Heights of women: Sometimes it helps to standardize measurement.

Suppose that the mean height of all women in the USA is $\mu = 65$ inches = 165.1 cm.

Suppose that the standard deviation of the height of all women in the USA is $\sigma = 2.5$ inches = 6.35 cm.

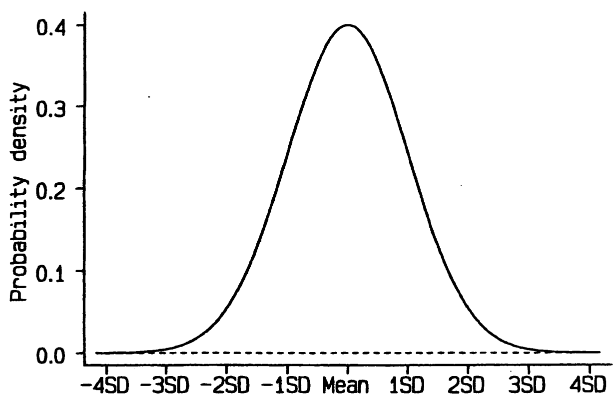
Now consider the particular height 70 inches = 177.8 cm.

Let $Z_1 = (70 \text{ in} - 65 \text{ in})/(2.5 \text{ in}) = 2$.

Let $Z_2 = (177.8 \text{ cm} - 165.1 \text{ cm})/(6.35 \text{ cm}) = 2$.

So, on the SD scale, the particular height 70 inches = 177.8 cm, becomes 2 SD regardless of the units.

The Normal distribution



Critical values of the Normal (= Gaussian) distribution.

Suppose that Z has the Normal (= Gaussian) distribution.

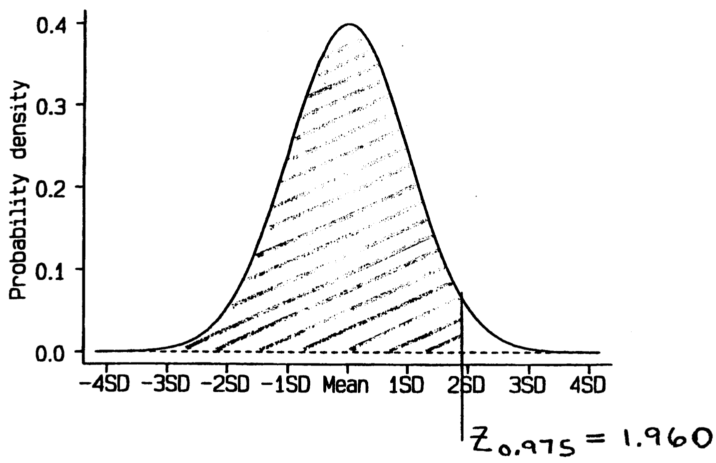
One can ask, what is the probability that Z is less than some critical value, some number of SDs.

$$\text{Prob}(Z < Z_u) = \text{Prob}(Z \leq Z_u) = u.$$

For example,

$\text{Prob}(Z < Z_{0.8} = 0.841) = 0.8.$
$\text{Prob}(Z < Z_{0.9} = 1.282) = 0.9.$
$\text{Prob}(Z < Z_{0.95} = 1.645) = 0.95.$
$\text{Prob}(Z < Z_{0.975} = 1.960) = 0.975.$
$\text{Prob}(Z < Z_{0.99} = 2.326) = 0.99.$
$\text{Prob}(Z < Z_{0.995} = 2.576) = 0.995.$

The Normal distribution



The Normal distribution is symmetric about zero. Thus, $Z_u = -Z_{1-u}$.

For example,

$\text{Prob}(Z < Z_{0.2} = \text{-}0.841) = 0.2.$

$\text{Prob}(Z < Z_{0.1} = \text{-}1.282) = 0.1.$

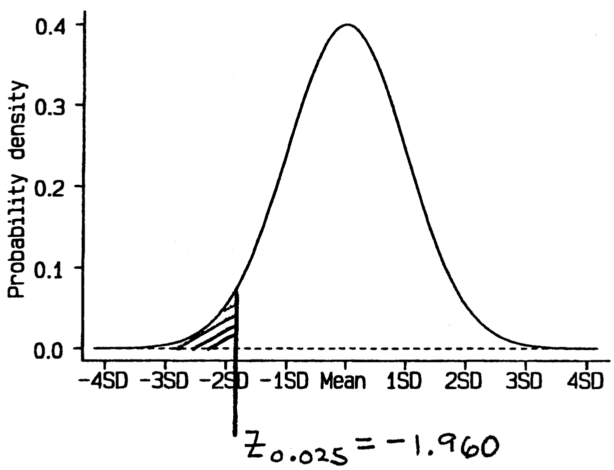
$\text{Prob}(Z < Z_{0.05} = \text{-}1.645) = 0.05.$

$\text{Prob}(Z < Z_{0.025} = \text{-}1.960) = 0.025.$

$\text{Prob}(Z < Z_{0.01} = \text{-}2.326) = 0.01.$

$\text{Prob}(Z < Z_{0.005} = \text{-}2.576) = 0.005.$

The Normal distribution



From these results, one can ask, what is the probability that Z is between two critical values.

$\text{Prob}(Z_u < Z < Z_{1-u}) = 1 - 2u.$

$\text{Prob}(Z_{u/2} < Z < Z_{1-u/2}) = 1 - u.$

For example:

$\text{Prob}(-0.841 = Z_{0.2} < Z < Z_{0.8} = 0.841) = 0.6.$

$\text{Prob}(-1.282 = Z_{0.1} < Z < Z_{0.9} = 1.282) = 0.8.$

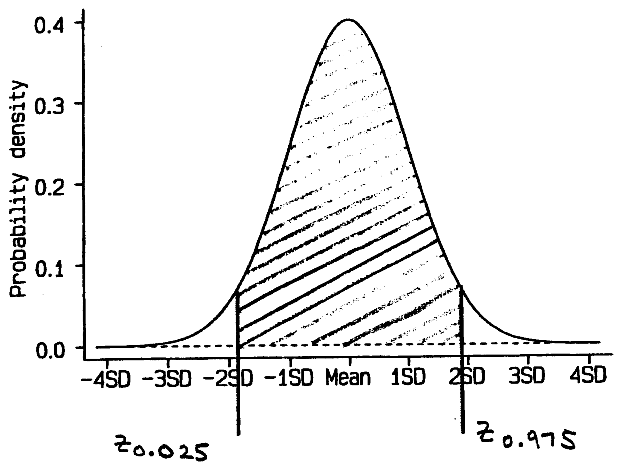
$\text{Prob}(-1.645 = Z_{0.05} < Z < Z_{0.95} = 1.645) = 0.9.$

$\text{Prob}(-1.960 = Z_{0.025} < Z < Z_{0.975} = 1.960) = 0.95.$

$\text{Prob}(-2.326 = Z_{0.01} < Z < Z_{0.99} = 2.326) = 0.98.$

$\text{Prob}(-2.576 = Z_{0.005} < Z < Z_{0.995} = 2.576) = 0.99.$

The Normal distribution



Conversely, one can ask, what is the probability that Z is outside of two critical values.

$\text{Prob}(Z < Z_u \text{ or } Z > Z_{1-u}) = 2u.$

$\text{Prob}(Z < Z_{u/2} \text{ or } Z > Z_{1-u./2}) = u.$

For example:

$\text{Prob}(Z < Z_{0.2} = \text{-}0.841 \text{ or } Z > Z_{0.8} = 0.841) = 0.4.$

$\text{Prob}(Z < Z_{0.1} = \text{-}1.282 \text{ or } Z > Z_{0.9} = 1.282) = 0.2.$

$\text{Prob}(Z < Z_{0.05} = \text{-}1.645 \text{ or } Z > Z_{0.95} = 1.645) = 0.1.$

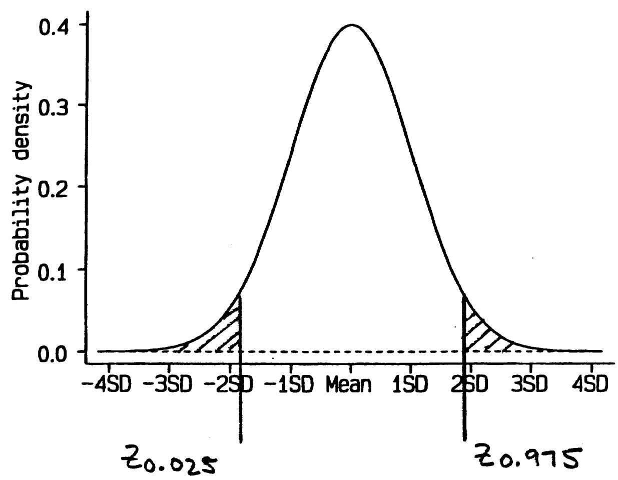
$\text{Prob}(Z < Z_{0.025} = \text{-}1.960 \text{ or } Z > Z_{0.975} = 1.960) = 0.05.$

$\text{Prob}(Z < Z_{0.01} = \text{-}2.326 \text{ or } Z > Z_{0.99} = 2.326) = 0.02.$

$\text{Prob}(Z < Z_{0.005} = \text{-}2.576 \text{ or } Z > Z_{0.995} = 2.576) = 0.01.$

This result is parallel to hypothesis testing and confidence interval construction.

The Normal distribution



Example: Heights of women, continued.

Suppose that the mean height of all women in the USA is μ = 65 inches = 165.1 cm and the SD is σ = 2.5 inches = 6.35 cm.

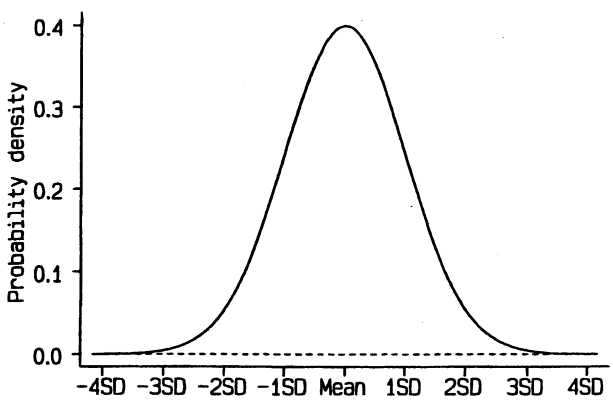
One can ask, what is the probability that a randomly selected height is less than 70 inches = 177.8 cm = 2 SD (treating height as a continuous variable).

Prob(Z < 2) = 0.977.

Thus, 70 inches = 177.8 cm = 2 SD corresponds to the 97.7th percentile.

Note that 86.3% of observations fall within 1 SD of the mean, 95.4% of observations fall within 2 SDs of the mean, and 99.73% of observations fall within 3 SDs of the mean.

Range	Probability within range	Probability outside range
Mean ± 1SD	68.3%	31.7%
Mean ± 2SD	95.4%	4.6%
Mean ± 3SD	99.73%	0.27%



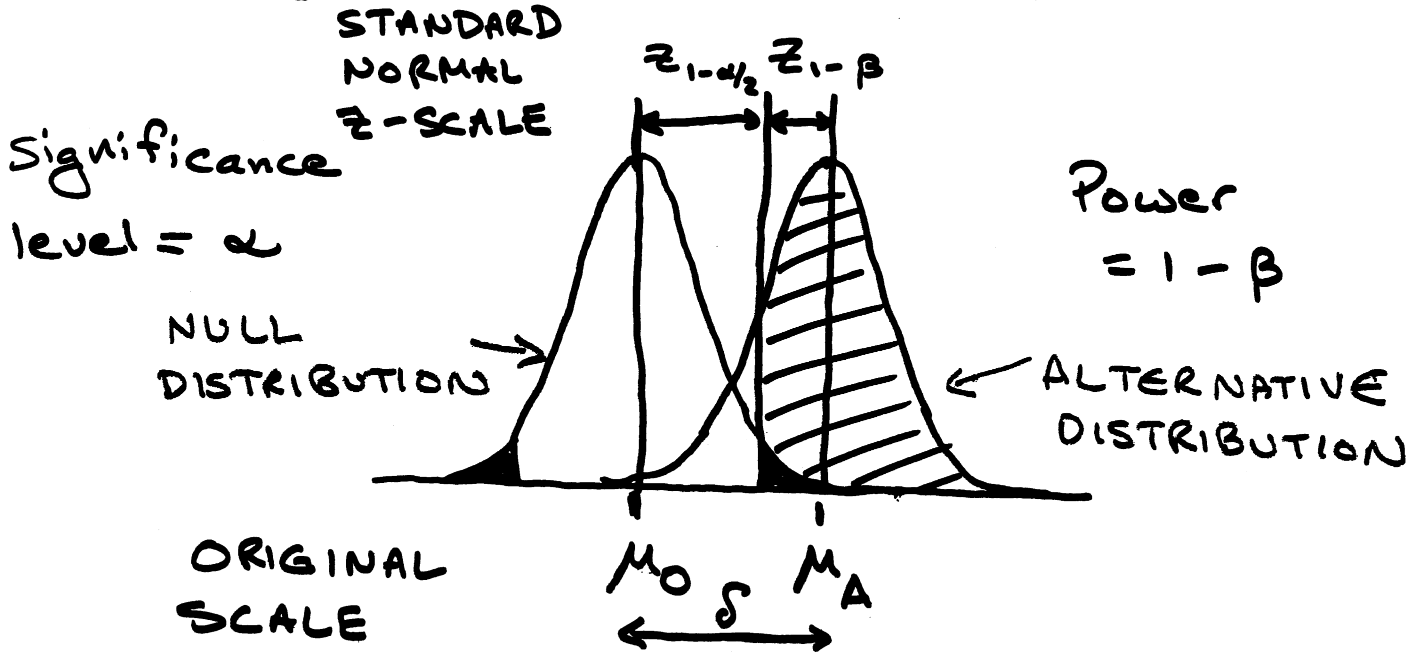
The basic idea behind sample size calculations.

- Set the Type I error rate, α. Then find $Z_{1-\alpha/2}$. (two sided.)
- Set the Type II error rate, β. Then find $Z_{1-\beta}$.
- Choose the hypotheses of interest.
  H$_0$: $\mu_1 = \mu_0$ vs. H$_A$: $\mu_1 = \mu_A \neq \mu_0$. (two-sided.)
- Let $\delta = \mu_1 - \mu_0$ denote the difference to be detected.
- Let $\sigma/\sqrt{n}$ denote the standard error (SE) of some estimate of δ.

The trick of sample size estimation is to equate $Z_{1-\alpha/2} + Z_{1-\beta}$ with $\delta/(\sigma/\sqrt{n}) = \sqrt{n}\Delta$ with $\Delta = \delta/\sigma$.

$$Z_{1-\alpha/2} + Z_{1-\beta} = \sqrt{n}\Delta$$
$$(Z_{1-\alpha/2} + Z_{1-\beta})^2 = n\Delta^2$$
$$n = (Z_{1-\alpha/2} + Z_{1-\beta})^2 / \Delta^2.$$

A schematic picture:

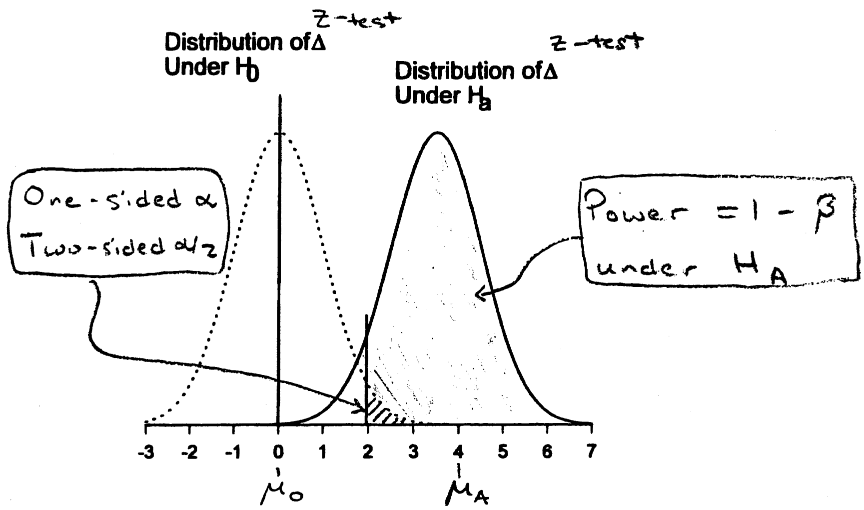


[2] Basic concepts in sample size and power calculation.

The power of a test.

- Recall that the *power of a test* is the probability of rejecting the null hypothesis (H_0) when it is false.
- Power = 1 – P(Type II error) = 1 - β.
- Power is computed for a *particular* value of the alternative hypothesis (H_A).

Figure from Piantadosi (1997, Figure 7.1, p. 163):



[redacted]re 7.1 Distributions of an estimator under the null and alternative hypothe-[redacted]Vertical lines are drawn at $\Delta = 0$ and $c = 1.96$ as explained in the text.

[4.3] An alternative approach: Two stage designs.

Two-stage designs are useful for Phase II (safety/efficacy) clinical trails. Two-stage designs are optimal because under the null hypothesis they have the smallest *expected* sample size.

- Choose α and β.
- Choose π_0, the maximum clinically uninteresting success rate, and π_A, the minimum clinically meaningful success rate.
- In stage 1 enroll n_1 patients. If the number of responses is less than or equal to r_1, abandon the new treatment. The design does not permit stopping early for efficacy.
- In stage 2 enroll up to n patients sequentially (including the original n_1 in stage 1). If the number of responses is less than or equal to r, abandon the new treatment. If the number of responses is greater than r, accept the treatment for further study.

Table from Piantadosi (1997, Table 7.4, p. 160):

Table 7.4 Optimal Two-Stage Designs for SE (Phase II) Trials for $p_1 - p_0 = 0.20$ and $\alpha = 0.05$

$\delta = \pi_0 - \pi_A = 0.20$

Power = $1-\beta$

π_0 p_0	π_A p_1	β	r_1	n_1	r	n	$E\{n \mid p_0\}$*
.05	.25	.2	0	9	2	17	12
		.1	0	9	3	30	17
.10	.30	.2	1	10	5	29	15
		.1	2	18	6	35	23
.20	.40	.2	3	13	12	43	21
		.1	4	19	15	54	30
.30	.50	.2	5	15	18	46	24
		.1	8	24	24	63	35
.40	.60	.2	7	16	23	46	25
		.1	11	25	32	66	36
.50	.70	.2	8	15	26	43	24
		.1	13	24	36	61	34
.60	.80	.2	7	11	30	43	21
		.1	12	19	37	53	30
.70	.90	.2	4	6	22	27	15
		.1	11	15	29	36	21

* Gives the expected sample size when the true response rate is p_0. π_0